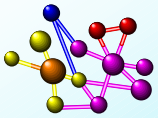
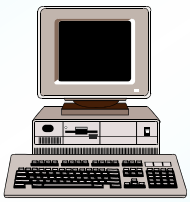
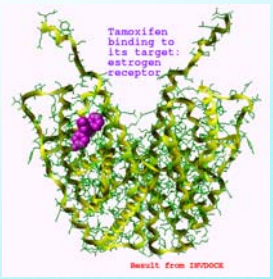
INVDOCK: A Method and Software for Computer Automated Prediction of Protein Targets of Small Molecules

Chen Yu Zong

- Purpose.
- Background Info.
- Description of Technology.
- Performance Analysis







Purpose

- To provide a new method for low-cost and high-speed prediction of protein and nucleic acid targets of a small molecule.
- Potential applications:

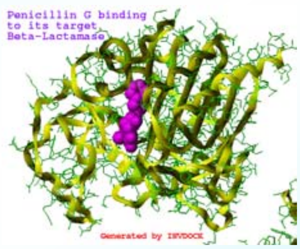
1. Identification of unknown and secondary therapeutic targets of drugs, drug leads, drug candidates, natural products, etc.
2. Prediction of drug targets related to side effect and toxicity (drug safety evaluation).
3. Prediction of targets related to drug ADME (pharmacokinetics analysis).
4. Identification of unknown receptors of a ligand (pathway analysis).

Background Info

Why study protein targets of a molecule?

Therapeutic Targets

Nature 396, 15 (1998)



now comes from electron microscopy on decagonal $Al_{72}Ni_{20}Co_{8}$. Steinhardt and colleagues[2] present micrographs that demonstrate the existence of Gummelt's decagonal coverage in nature. They used the high-angle annular dark-field technique[11]. In contrast to classical high-resolution transmission electron microscopy, where contrast interpretation requires a careful quantum-mechanical simulation of the dynamic imaging process (often with rather ambiguous results), the dark-field technique is based on high-angle scattering of a fine beam of electrons at atom cores in a scanning transmission electron microscope. In this case the imaging process is essentially kinematic in nature and allows a direct, intuitive image interpretation in terms of atomic structure. This makes the new results superior to those of previous studies in which atomic matching of classical high-resolution images to model lattices has been claimed.

Steinhardt and colleagues call the decorated decagon a 'quasi-unit cell' to emphasize similarities to ordinary crystal lattices. Whether this is a happy choice is questionable as clusters often also play a role in crystals; an example is the Frank–Kasper phases (where clusters occur which are quite similar to those found in quasicrystals). But these clusters never coincide with the crystal unit cells.

That clusters are important for quasicrystal formation was noticed quite early on[13,14]. In particular, the model of icosahedral Al–Pd–Mn, the quasicrystal of highest structural quality available today, is based on a 52-atom icosahedral cluster[14]. In this case it could even be shown that it owes its particular stability to a low electron energy[15]. In its present form, however, this model does not explicitly involve cluster overlap as a construction principle, but employs 'connecting units' which are interpenetrating pieces of the elementary cluster. Indeed, the new 'covering' model has still to be tested for three-dimensionally quasiperiodic quasicrystals. But the idea that the particular structure of quasicrystals is the consequence of maximizing the density of a low-energy atom cluster and (what appears to be particularly important) the experimental observation of the Gummelt coverage, make it worth while revising current models to incorporate this new view of quasicrystal formation. □

Knut W. Urban is at the Institute of Solid State Physics, Helmholtz Research Centre Jülich, D-52425 Jülich, Germany.
e-mail: k.urban@fz-juelich.de

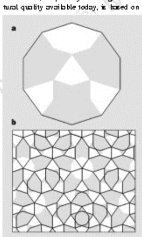


Figure 2 Gummelt's planar quasiperiodic coverage. a, Gummelt's decorated decagon. The matching rule demands that two decagons may overlap only if shaded areas overlap. b, The resulting quasiperiodic decagonal coverage of the plane.

Inflammation

A new target for aspirin

Edward A. O'Neill

"Take two aspirin and call me in the morning" seems like a tired cliché. Nevertheless, for many common physical ailments, low doses of aspirin (acetylsalicylic acid) will make you feel better, and high doses of the drug give sustained relief from the symptoms of chronic inflammatory diseases such as rheumatoid arthritis[1]. Aspirin irreversibly inhibits the cyclooxygenases — enzymes that control the production of prostaglandins (small molecules that induce the pain and fever associated with infection or trauma[2]). However, the difference in the clinical activities of aspirin at low and high doses has led to speculation that not all the benefits of aspirin derive from inhibition of cyclooxygenases. So what might be the target for high doses of the drug? On page 77 of this issue, Yin et al.[3] report that high concentrations of aspirin inhibit the recently discovered enzyme IκB kinase-β (IKK-β), and they propose that this effect partially explains the clinical efficacy of high-dose aspirin.

The IKK enzymes (α and β) catalyse the transfer of phosphate moieties from ATP to IκB (reviewed in ref. 4). Phosphorylation leads to the degradation of IκB and release of NF-κB, a transcription factor that is inhibited by IκB. On release, NF-κB rapidly moves to the nucleus where it binds specific DNA sequences, promoting the transcription of genes that influence defence mechanisms such as inflammatory and immune responses. Thus, if IKK-β is inhibited by aspirin, nuclear localization of NF-κB and subsequent transcription should be blocked. This has, in fact, been observed with high concentrations of aspirin[5].

Yin et al.[3] now report that high concentrations of aspirin (IC_{50} approximately 50 μM) are required to inhibit IKK-β. They also show that several other kinases, including the homologous and functionally related IKK-α, are not affected by aspirin. Can inhibition of IKK-β and cyclooxygenases explain all the pharmacological effects of aspirin? Let's hope not. The problem with using high doses of aspirin is that it has virtually no therapeutic window — that is, the dose at which aspirin gives relief from chronic rheumatic disease is very close to the dose that generates side-effects, including headaches, dizziness and tinnitus[6]. But, based on the work of Yin and colleagues, we can develop a hypothesis wherein IKK-β is only one of several targets for high-dose aspirin, the unwanted side-effects resulting from activity against other targets.

The anti-inflammatory efficacy of aspirin may, therefore, be limited by two separable factors. First, aspirin weakly inhibits the therapeutic target (IKK-β). Second, it is a similarly weak inhibitor of other, as-yet-unidentified targets, yielding various toxicities. There are many reports of aspirin affecting a variety of biological systems that could be independent of IKK-β inhibition. Most interesting in light of the paper by Yin et al. are reports that high concentrations of aspirin inhibit the activation of the Jun-N-terminal kinases (JNKs)[7]. The JNK family of protein kinases mediates phosphorylation of the transcription factor c-Jun, stimulating its ability to promote transcription[8]. Because there is no evidence that IKK-β is involved in activating JNK, these results imply that at least one other target of high-dose aspirin exists.

Although there is a broad consensus that aspirin relieves many of the symptoms asso-

Background Info

Why study protein targets?

Prediction of side effect, toxicity, pharmacokinetics and pharmacogenetics

Annu. Rev. Pharmacol. Toxicol. 2000. 40:353–88

MITOCHONDRIAL TARGETS OF DRUG TOXICITY

K. B. Wallace and A. A. Starkov
Department of Biochemistry and Molecular Biology, University of Minnesota School of Medicine, Duluth, Minnesota 55812; e-mail: kwallace@d.umn.edu, astarkov@d.umn.edu

Key Words oxidative phosphorylation, uncouplers, bioenergetics, permeability transition, redox cycling

■ **Abstract** Mitochondria have long been recognized as the generators of energy for the cell. Like any other power source, however, mitochondria are highly vulnerable to inhibition or uncoupling of the energy harnessing process and run a high risk for catastrophic damage to the cell. The exquisite structural and functional characteristics of mitochondria provide a number of primary targets for xenobiotic-induced bioenergetic failure. They also provide opportunities for selective delivery of drugs to the mitochondrion. In light of the large number of natural, commercial, pharmaceutical, and environmental chemicals that manifest their toxicity by interfering with mitochondrial bioenergetics, it is important to understand the underlying mechanisms. The significance is further underscored by the recent identification of bioenergetic control points for cell replication and differentiation and the realization that mitochondria play a determinant role in cell signaling and apoptotic modes of cell death.

Annu. Rev. Pharmacol Toxicol 2000, 40:353-388
1997, 37:269-296

Annu. Rev. Pharmacol. Toxicol. 1997. 37:269–96

MOLECULAR MECHANISMS OF GENETIC POLYMORPHISMS OF DRUG METABOLISM

Urs A. Meyer
Biozentrum of the University of Basel, CH-4056 Basel, Switzerland

Ulrich M. Zanger
Dr. Margarete Fischer-Bosch-Institute for Clinical Pharmacology, D-70376 Stuttgart, Germany

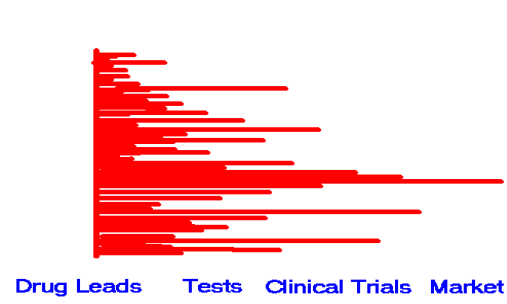
KEY WORDS: Genetic polymorphism, CYP2D6, CYP2C19, N-acetyltransferases, drug metabolism, pharmacogenetics

ABSTRACT

One of the major causes of interindividual variation of drug effects is genetic variation of drug metabolism. Genetic polymorphisms of drug-metabolizing enzymes give rise to distinct subgroups in the population that differ in their ability to perform certain drug biotransformation reactions. Polymorphisms are generated by mutations in the genes for these enzymes, which cause decreased, increased, or absent enzyme expression or activity by multiple molecular mechanisms. Moreover, the variant alleles exist in the population at relatively high frequency. Genetic polymorphisms have been described for most drug metabolizing enzymes. The molecular mechanisms of three polymorphisms are reviewed here.

Background Info

Detection of side effect and toxicity in early stages of drug discovery



Drug Candidates
in Different Stages of Development
Majority of Them Fail to Reach Market
Clin Pharmacol Ther. 1991; 50:471

Significance:

- Most drug candidates fail to reach market (> 99%)
- Side effect and toxicity is an important reason (in 30-40% cases).
- Large portion of money ($350 million per drug) and time (6-12 years for a drug) has been wasted on failed drugs.

Drug Discov Today 1997; 2:72

Background Info

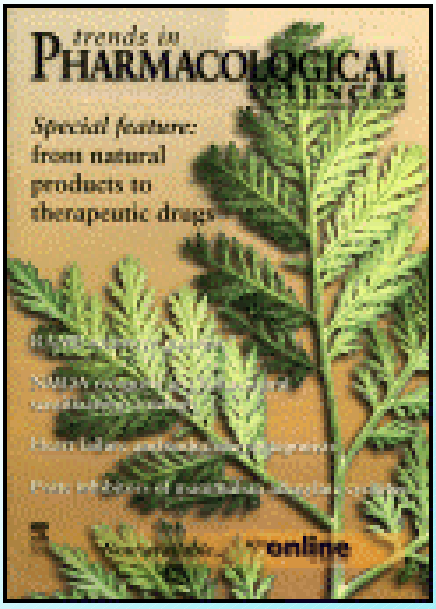
Why study protein targets of a molecule?

Drugs from Natural Products

From natural products
to therapeutic drugs

TIPS, May 1999, 20:190

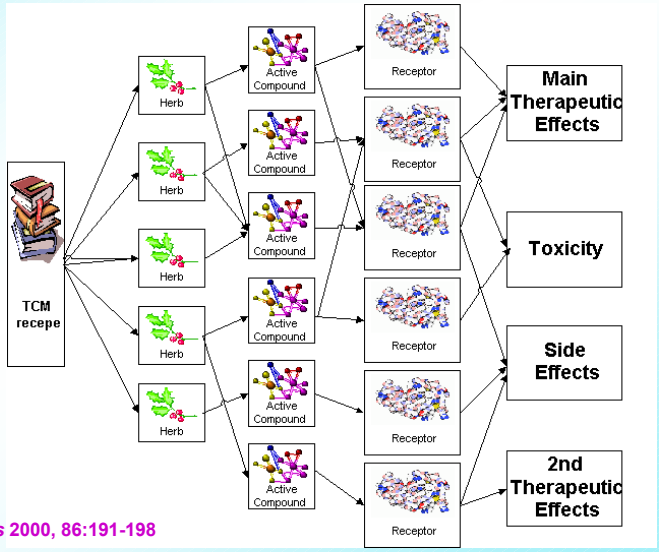
- Screening of bioactive compounds
- Molecular mechanism
- Further development



Background Info

Drug targets and new drug discovery
Drugs from Traditional Medicine

- Mixture of multiple herbs etc.
- Maintaining and restoring balance.
- Mutual accentuation, mutual enhancement, mutual counteraction, mutual suppression, mutual antagonism, mutual incompatibility
- Multiple targets: therapeutic effects, symptom treatment, toxicity modulation, drug delivery, harmonization



Pharmacology & Therapeutics 2000, 86:191-198

Background Info

The need for a new target-prediction method:

- Existing experimental methods costly and time-consuming.
- Limited resources: difficulty in compound synthesis and bioassay.
- Existing computer methods not designed for target searching. Inability in cavity identification and in docking to large cavities.

Background Info

Feasibility:

Protein sampling:

- Database: >15,000 3D structures in PDB.
- Protein diversity: 17% in PDB with unique sequence.
- Development of structural genomics: 10,000 unique proteins within 5 years.

Ann. Rev. Biophys. Biomol. Struct. 1996; 25:113
Nature Struct. Biol. 1998; 5:1029

Method Accuracy:

- Ligand-protein docking algorithms capable of finding binding conformations.

Proteins. 1999; 36:1

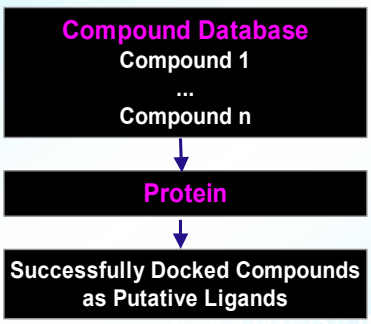
Computer capability:

- Increasing performance (docking of 100,000 compounds in days).
- Decreasing cost (Linux PC, Multi-processor Machine)

Technology Description
Strategy

Existing Methods:

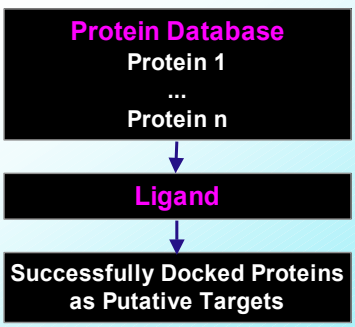
Given a Protein,
Find Putative Binding Ligands
From a Chemical Database



Science 1992;257: 1078

New Method:

Given a Ligand,
Find Putative Protein Targets
From a Protein Database



Proteins 2001;43: 217

Technology Description

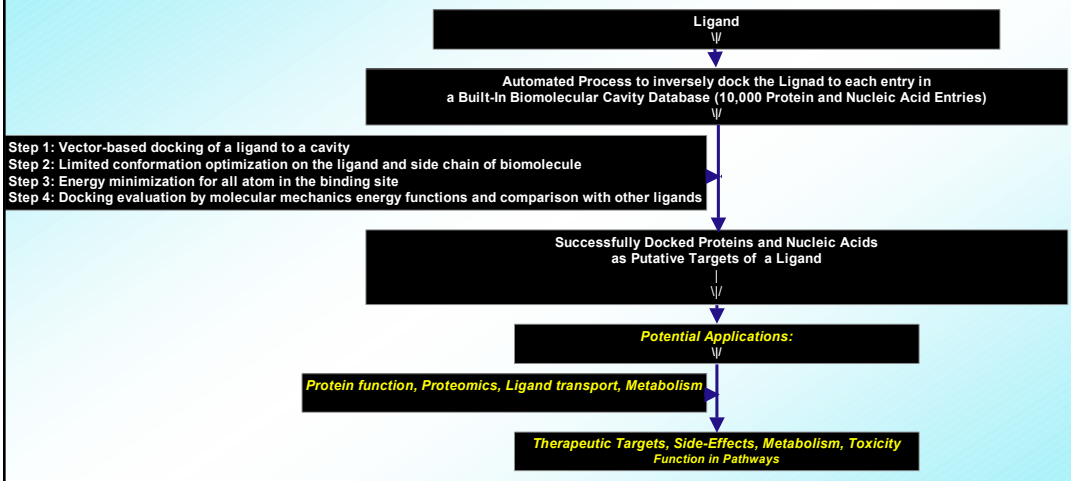
Key technology:

- Ligand-protein inverse docking strategy for target identification through structure database search.
- Flexible ligand-protein docking algorithm with no restriction on cavity size and no knowledge about specific binding region within a cavity.
- Method for automated detection of all cavities in a protein or a nucleic acid.
- Development of a biomolecular cavity database for all protein and nucleic acid entries in PDB.

Proteins 2001;43: 217

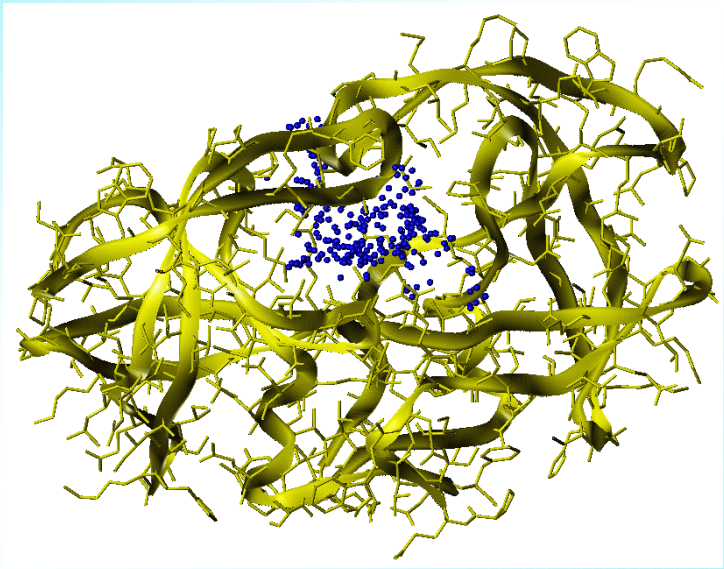
Technology Description

INVDOCK procedure:



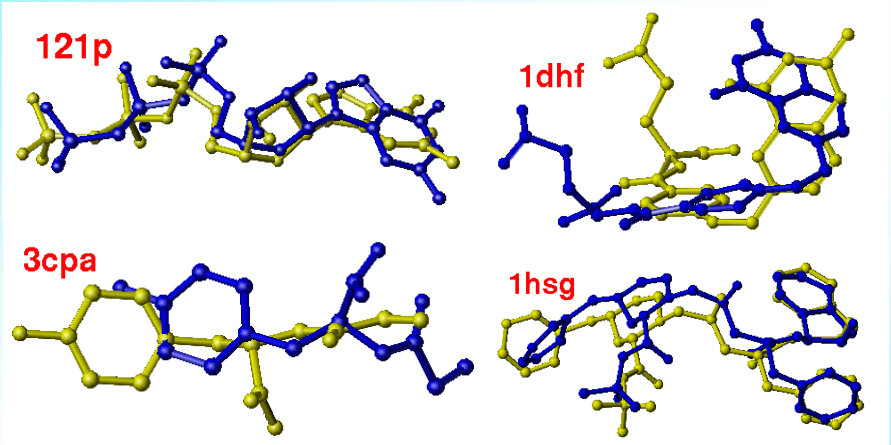
Technology Description

INVDOCK Cavity Models



HIV-1 Protease

INVDOCK Performance Analysis



The docked (blue) and crystal (yellow) structure of ligands in some PDB ligand-protein complexes. The PDB Id of each structure is shown.

INVDOCK Performance Analysis

Molecule	Docked Protein	PDB Id	RMSD	Description of Docking Quality	Energy
Indinavir	HIV-1 Protease	1hsg	1.38	Match	-70.25
Xk263 Of Dupont Merck	HIV-1 Protease	1hvr	2.05	Match	-58.07
Vac	HIV-1 Protease	4phv	0.80	Match	-88.46
Folate	Dihydrofolate Reductase	1dhf	2.41	One end match, the other in slightly different orientation	-63.02
5-Deazafolate	Dihydrofolate Reductase	2dhf	1.48	Match	-65.49
Estrogen	Estrogen Receptor	1a52	1.30	Match	-45.86
4-Hydroxytamoxifen	Estrogen Receptor	3ert	0.97	Match	-55.15
Guanosine-5'-[B,G-Methylene] Triphosphate	H-Ras P21	121p	0.94	Match	-80.20
Glycyl-*L-Tyrosine	Carboxypeptidase A α	3cpa	2.19	Match	-44.84

Proteins 2001;43: 217

INVDOCK Performance Analysis

Compound	Potential Targets Identified	Experimentally Confirmed	Experimentally Implicated
4H-Tamoxifen	17	4	4
Aspirin	52	4	16
Vitamin C	46	4	9
Vitamin E	26	2	11

Proteins 2001;43: 217

INVDOCK Performance Analysis

Compound	Number of experimentally confirmed or implicated toxicity targets	Number of toxicity targets predicted by INVDOCK	Number of toxicity targets missed by INVDOCK	Number of toxicity targets without 3D structure or involving covalent bond	Number of INVDOCK predicted toxicity targets without experimental finding
Aspirin	15	9	2	4	2
Gentamicin	17	5	2	10	2
Ibuprofen	5	3	0	2	2
Indinavir	6	4	0	2	2
Neomycin	14	7	1	6	6
Penicillin G	7	6	0	1	8
Tamoxifen	2	2	0	0	4
Vitamin C	2	2	0	0	3
Total	68	38	5	25	29

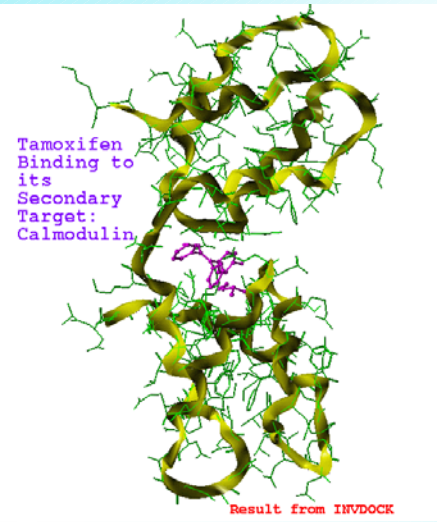
J. Mol. Graph. Mod 2001;20: 199

INVDOCK Performance Analysis

Natural Product Drug from Chinese Medicinal Plants	Number of Identified Therapeutic Targets	Number of Confirmed or Implicated Therapeutic Targets by experiment	Number of Identified Toxicity/Side effect Targets	Number of Confirmed or Implicated Toxicity/Side Effect Targets by experiment
Acronycine	3	1	4	-
Allicin	5	2	1	1
Baicalin	14	4	6	-
Catechin	17	12	5	-
Camptothecine	9	6	3	2
Dicoumarin	7	1	3	1
Emodin	6	3	5	1
Genistin	22	7	12	1

Chin. J. Med. Chem 2001;11: 145; *Am. J. Chin. Med.* 2002, in print

Conclusions



- Ligand-protein inverse docking is useful in probing potential targets of a molecule (~50% accuracy).
- Potential application in drug therapeutic target identification, safety evaluation, and pharmacokinetics prediction for drugs, drug leads, and natural products.
- Application potential increases along with advances in structural and functional genomics.